Check for
updates

PROJECT NOTES

A corpus of Schlieren photography of speech production: potential methodology to study aerodynamics of labial, nasal and vocalic processes

Fabian Tomaschek[1] · Denis Arnold[2] · Konstantin Sering[1] · Friedolin Strauss[3]

Abstract This report presents a corpus of articulations recorded with Schlieren photography, a recording technique to visualize aeroflow dynamics for two purposes. First, as a means to investigate aerodynamic processes during speech production without any obstruction of the lips and the nose. Second, to provide material for lecturers of phonetics to illustrates these aerodynamic processes. Speech production was recorded with 10 kHz frame rate for statistical video analyses. Downsampled videos (500 Hz) were uploald to a youtube channel for illustrative purposes. Preliminary analyses demonstrate potential in applying Schlieren photography in research.

Keywords Articulation · Speech production · Schlieren photography · Aerodynamics

1 Introduction

Phonetic scientists have created and applied a set of methods and techniques to investigate the physical characteristics and the temporal dynamics of oral and nasal processes during speech production. Kinematic processes have been studied with electromagnetic articulography (e.g. Hoole et al., 1994; Mooshammer & Fuchs, 2002; Tiede et al., 2001; Tomaschek et al., 2018), Ultrasound (e.g. Davidson, 2006;

✉ Fabian Tomaschek
    fabian.tomaschek@uni-tuebingen.de

1   Department of General Linguistics, University of Tübingen, Tübingen, Germany

2   Leibniz-Institut für Deutsche Sprache, Digitale Sprachwissenschaft, Mannheim, Germany

3   Deutsches Zentrum für Luft- und Raumfahrt DLR e.V., Institut für Raumfahrtantriebe, Hardthausen, Germany

Zharkova et al., 2012; Wrench & Scobbie, 2011) and even magnetic resonance imaging (MRI) (e.g. Mathiak et al., 2000; Uecker et al., 2010). Airflow processes have been studied by means of pressure transducers (e.g. Basset et al., 2001; Petrone et al., 2017; Hertegård & Gauffin, 1992) or a Rothenberg mask (e.g. Quigley et al., 1964; Warren, 1967). All of these techniques come with their specific advantages, such as high spatial and high temporal resolution. However, when it comes to investigating air flow, there are crucial disadvantages in the available techniques. For example, intra-oral and intra-nasal pressure is measured by inserting tubes into the nose. Not only comes this technique with potentially high discomfort for participants, there is a high risk that saliva blocks the tube, reducing the amount of recording time. Even though Rothenberg masks avoid the problem of discomfort and short recording times, they systematically change the characteristics of the acoustic signal and block the jaw, leading to additional changes in speech production. In the current paper, we present a corpus of Schlieren photography, a technique that avoids these disadvantages. We also provide technical instructions how the necessary recording technology can be purchased and how this set up, under the guidance of a skilled physicist, can be built.

Schlieren photography has already been proposed 40 years ago as a tool to investigate speech (e.g. Davies, 1979, 1981). However, only one proper study using Schlieren photography has so far been published. Rowell et al. (2016) used the technique to investigate the production of french oral and nasal vowels. Krane and Gary (2004) published an abstract on investigating fricatives using Schlieren photography, but the actual paper has never been published.

Schlieren photography allows to record the dynamics of oral and nasal airflow without obstruction of the labial articulatory processes. The material was therefore recorded for two purposes. First, as a means to investigate aerodynamic processes during speech production without any obstruction of the lips. Second, to provide material for lecturers of phonetics to illustrates these aerodynamic processes. In the remainder of this paper, we describe Schlieren photography and discuss its limitations. We report the linguistic material chosen for the corpus and the recording and preprocessing of the data. We close the presentation with an illustrative description of aerodynamics in various combinations of manner and place of articulation of consonants, in addition to a preliminary analysis with the present material.

1.1 Description of Schlieren photography

Schlieren photography visualizes the small disturbances and refractive differences within the background of inhomogeneous transparent media caused e.g. by a density or pressure gradient (Settles, 2001). Those refractive differences bend light beams by definition into non-normal directions (z-axes). The basic principle relies on a wide beam of parallel light emitted by a point light source and paralleled by lenses. This parallel light beam consists of an infinitesimal number of smaller parallel light rays which are focused on a knife edge and projected onto a screen or a camera sensor. The knife edge is used to increase the sensitivity of the system but also to block the image of the light source from being projected to the screen/camera

sensor. A Schlieren object (in this case density gradients) in the test area (between the two parallelization lenses) will bend the light rays e.g. in y-direction. A part of these bent light rays will then be blocked by the knife edge located in the focal point. The Schlieren object (in this case density gradients) are then displayed as dark shadows against a bright background.

1.2 Limitations of Schlieren photography

If readers of this papers would like to build a Schlieren set up on their own, we would like to inform them about certain limintations of Schlieren photography. In the publication at hand the needed density gradients were generated by the flow velocity of the air exiting the speaker's mouth and by the higher temperature of the air leaving the oral cavity. In order to increase this effect even more, the authors decided to heat up the test person's mouth by drinking hot beverages directly before a test run. See color plates in Settles (2001, p. 373) and Van Dyke (1997) for examples of exhalation flow and different Schlieren images.

It is important to note that changes in the refractive index can be caused by airflow or due to propagating waves through the air or even through both. Such changes can be caused e.g. by temperature differences and pressure differences or both at the same time as well as a gas of a different density being present in the flow (does not apply to our setup, where an exhalation air free-stream flow into ambient air was used). Unfortunately, the Schlieren photography cannot distinguish between the exact sources of change in refractive index (caused by density changes). In addition, the Schlieren method is also limited by its optical resolution. This maximum resolution is defined by the equipment applied (mainly the lenses, the camera and the mirror(s) involved) and the length of the optical path. The needed minimum resolution in order to resolve a given phenomenon as well as the maximum resolution with a given setup can be calculated with the formula given in Settles (2001). The application of this method is also limited to changes in the refractive index and optical transparent material. If there is no change in the refractive index or if this change is compensated before entering the camera lens system, nothing will be detected.

2 Description of the corpus

2.1 Material

Two speakers (authors FT & KS) uttered the disyllabic CVCV nonsense words in recordings 1 to 9. Author DA uttered recording 10.

(1) [fafa], [fifi], [fufu], [fõfõ]
(2) [mama], [mimi], [mumu], [mõmõ]
(3) [papa], [pipi], [pupu], [põpõ]
(4) [sasa], [sisi], [susu], [sõsõ]
(5) [zaza], [zizi], [zuzu], [zõzõ]

(6)   [papa], [pama], [mapa], [mala], [mama]
(7)   [pipi], [pimi], [mipi], [mili], [mimi]
(8)   [pupu], [pumu], [mupu], [mulu], [mumu]

(9)   [p͡fap͡fa], [ʃtaʃta], [ksaksa], [ʃpaʃpa]

(10)   [miːt͡səkat͡sə]

The speech material was designed to cover a maximum of potential questions. Sequences in 1 to 5 were designed such that the effects of an interaction between the different places and manners of articulation in interaction with vowel type on air flow dynamics could be investigated. Sequences in 6 to 8 were designed such that the interaction between a labial plosive consonant, a labial nasal consonant and vowel type could be investigated. The sequence in 9 was designed such that the effects of German affricates on air flow dynamics can be investigated. The final sequence is a German compound that features the affricate

.

2.2 Recordings and preprocessing

The experiments were conducted at the high-speed air breathing propulsion test facility M11.1 of DLR's Institute of Space Propulsion at the site of Lampold-shausen. This test bench is equipped with sensitive Schlieren systems as well as high speed cameras to investigate supersonic flows (see Strauss et al., 2016, 2017, for details). In the scientific field of aerospace applications with supersonic flows Schlieren photography is well-established and a standard non-intrusive measurement method. The Schlieren setup consisted of an optical rig, which was mounted in Toepler's z-type configuration (see Settles, 2001) for details on this configuration and Fig. 1). It featured two f/10 Schlieren mirrors with 150 mm diameter used to fold the beam path ("A" in Fig. 1). The setup had a total beam length of 7m and a measurement section length of 1m. A CREE® LED with a power of 3W and a colour temperature of 8000K provided the necessary light ("B" in Fig. 1). The light source was transferred into a point light source by a slit aperture in the beam path. A high speed black and white camera (PHOTRON® Fastcam SA1.1) with a Zeiss® Makro-Planar T* 2/100mm ZF lens was used to record the Schlieren images ("D" in Fig. 1) prepared by a Schlieren knife edge ("C" in Fig. 1).

A high-speed camera was chosen to detect high-frequency flow interference if present. The resolution of the PHOTRON® Fastcam SA1.1 camera was set to 384 × 380 pixel with a frame rate/temporal resolution of 10,000 Hz. The speakers seated themselves in the center of the optical set up (see Fig. 2). It was made sure that the shadow of the lower face (nose, mouth, and chin) was in the left third of the picture. The audio signal was simultaneously recorded with a nearly placed microphone (Audio-technica ATR1200®, cardioid recording plane) attached to portable Marantz® Professional PMD-670 recorder and manually synchronized

with the Schlieren photography in post-processing. A total of 14 files were recorded (317 GB of video material and 10 MB of audio).

2.3 Available data

Two sets of videos are available. For the first set, videos were downsampled to 500 Hz and uploaded to youtube[1], to allow lecturers an easy presentation of the material. Audio and downsampled videos can be retrieved from https://osf.io/z69ud/. The original videos with a frame rate of 10 kHz can be provided upon request. All material is published under (cc)-by license (https://creativecommons.org/licenses/by/4.0/deed.de).

3 Examples

In this section, we first provide a qualitative description of the Schlierenphotography. Subsequently, we demonstrate a statistical analysis of the data.

3.1 Qualitative description

We first contrast the articulation of the nasal [m], the fricatives [f] and [s], and the stops [p] and [t], all followed by the vowel [a]. Figures 3, 4, 5, 6 and 7 show examples of the recorded data. For a reference on manners and places of articulation see for instance Ladefoged and Maddieson (1996).

To articulate a nasal, the velum is lowered which allows the air to flow through the nasal passage. At the same time the oral cavity is closed at a point in front of the velum. In case of the [m] the lips are closed. In the first frame of Fig. 3, we can see that the lips are closed. The pattern of changes in the refractive index originates from the nose and widens over time (Fig. 3, frame 1). With the transition to the vowel, the changes in the refractive index originate from the mouth while the changes originating from the nose vanish (Fig. 3, frames 2 to 4).

Both, [f] and [s], are voiceless fricatives, i.e. a constriction is formed in the oral cavity causing turbulences in the airflow, while the vocal folds are not vibrating. The difference between [f] and [s] is the place of the constriction. To produce the [f], the constriction is formed between the lower lip and the upper teeth (Fig. 4, frame 1). The [s] is produced by bringing the tongue close to the roof of the mouth or in anatomical terms close to the palate. Comparing Figs. 4 and 5 (frames 2 to 4), we can observe different patterns in the refractive index of the air for both fricatives. While the changes appear to be smaller and more towards the front while uttering the [f], the changes seem to be stronger and more in a downward direction for the [s].

The last examples, the consonants [p] and [t], are both voiceless stops. During the articulation of a stop, the airflow is interrupted and rapidly released causing a burst. The interruption of the airflow is executed with both lips to produce a [p] (Fig. 6, frame 1) and with the tongue closing the oral cavity behind the teeth to articulate a

---

[1] https://www.youtube.com/playlist?list=PLLiPTCbSQtYm0Hi-tMXLqAiQGeUmtgsWg.
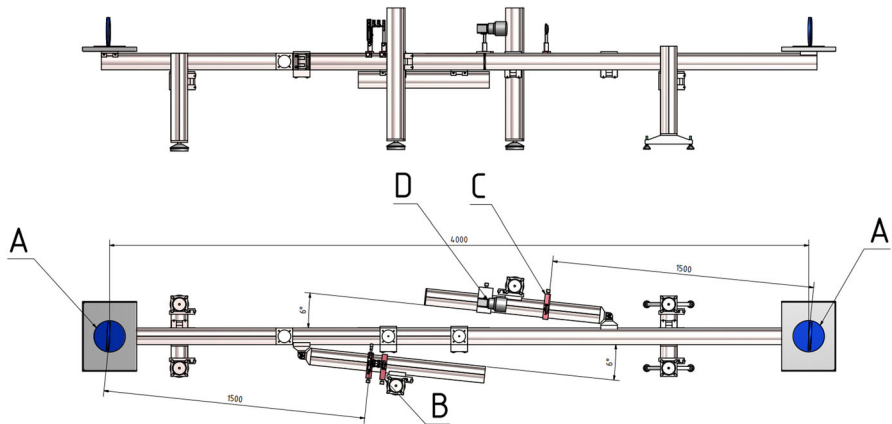
Fig. 1 Optical setup used for Schlieren photography at test bench M11.1 in Lampoldshausen. The speaker's head was positioned at the center between A and A
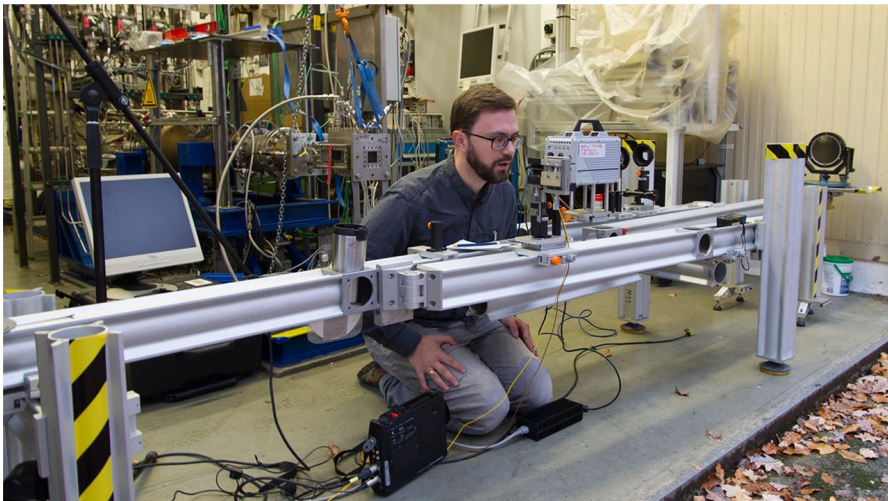

Fig. 2 Speaker's position in Schlieren setup at test bench M11.1 in Lampoldshausen

[t]. Comparing Figs. 6 and 7 (frames 2 to 4), we can observe again different patterns for the different places of articulation. The distortion pattern is more intense and more spreading downwards for [ta] in comparison to the pattern for [pa], which is spreading more towards the front.

Comparing the patterns of the consonants with a dental place of articulation [t] and [s] and consonants with a labial [p] and labiodental [f] place of articulation, we can observe that both dental consonants share a stronger, downward spreading pattern, which is probably caused by the airflow hitting the incisors where it is redirected in a downward direction.
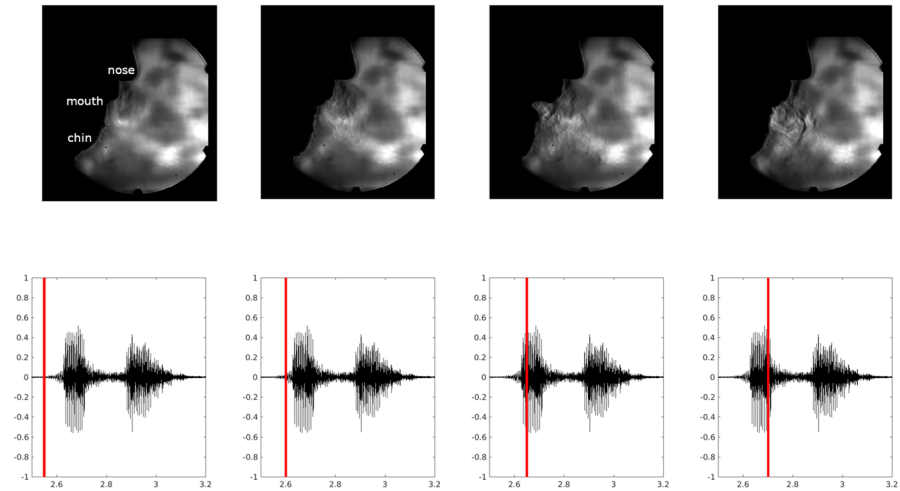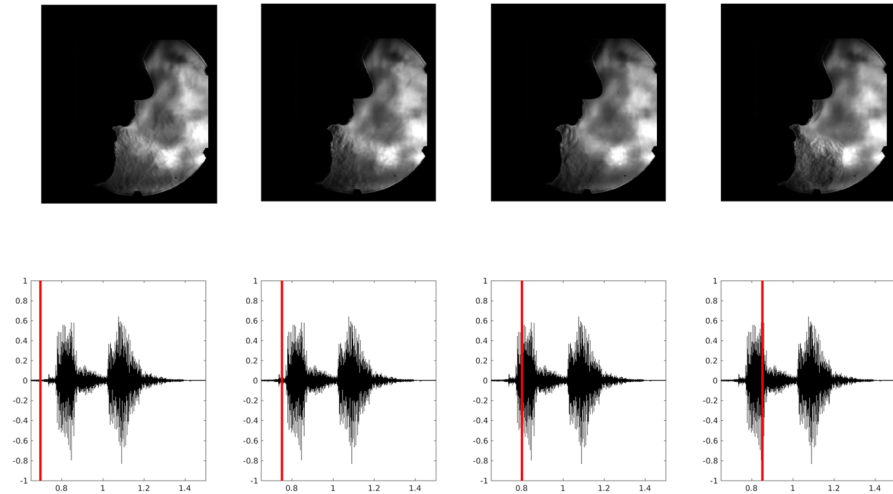
Fig. 3 Four frames of Schlieren photography (top panel) points in the articulation of [**mama**]. Bottom panel: oscillogram of the audio signal. The red line indicates the point in time. The first time point is before articulation of the consonant [m], the second time point during the articulation of [m], the last two time points are during the articulation of the vowel [a]
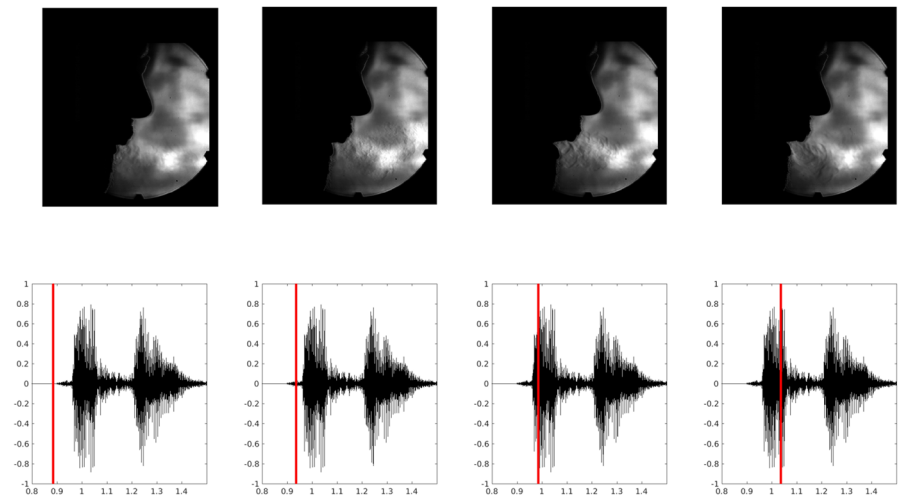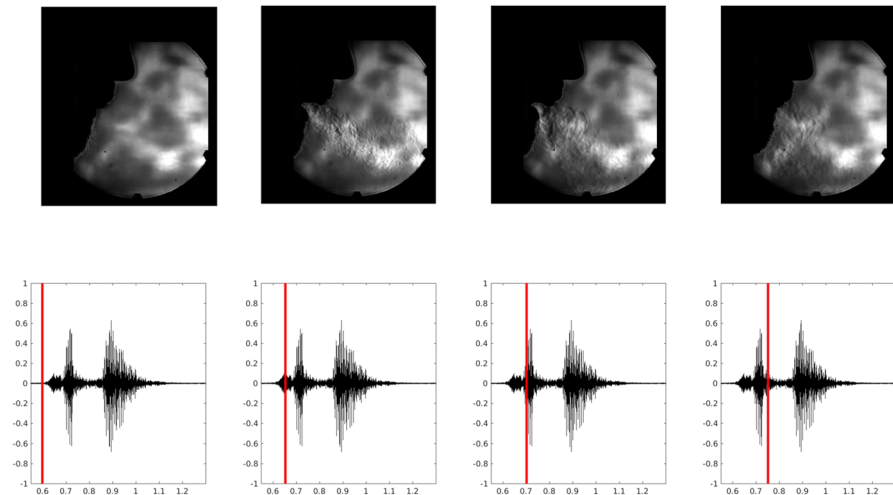

Fig. 4 Four frames of Schlieren photography (top panel) points in the articulation of [**fafa**]. Bottom panel: oscillogram of the audio signal. The red line indicates the point in time. The first time point is before articulation of the consonant [f], the second time point during the articulation of [f], the last two time points are during the articulation of the vowel [a]

We have shown that we find different patterns in the refractive index of the air for different manners of articulation and for the different places of articulation. The high temporal resolution of the video material allows a much more detailed observation of the changes in the patterns over time.

Fig. 5 Four frames of Schlieren photography (top panel) points in the articulation of [**sasa**]. Bottom panel: oscillogram of the audio signal. The red line indicates the point in time. The first time point is before articulation of the consonant [s], the second time point during the articulation of [s], the last two time points are during the articulation of the vowel [a]
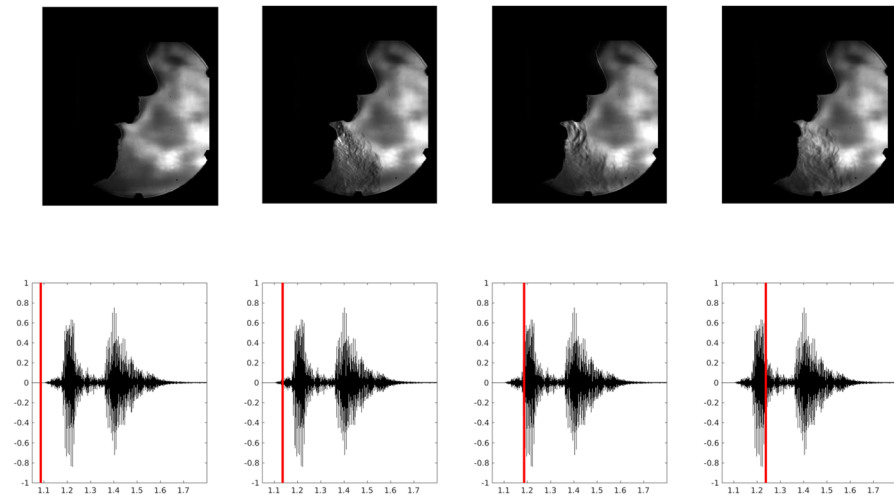

Fig. 6 Four frames of Schlieren photography (top panel) points in the articulation of [**papa**]. Bottom panel: oscillogram of the audio signal. The red line indicates the respective point in time. The first time point is before articulation of the consonant [p], the second time point during the articulation of [p], the last two time points are during the articulation of the vowel [a]

3.2 A qualitative approach to the analysis

In this section, we demonstrate a qualitative analysis of nasal airflow in the words 'papa', 'pama', 'mapa', 'mala' and 'mama'. In addition to nasal airflow during [m],

Fig. 7 Four frames of Schlieren photography (top panel) points in the articulation of [**tata**]. Bottom panel: oscillogram of the audio signal. The red line indicates the point in time. The first time point is before articulation of the consonant [t], the second time point during the articulation of [t], the last two time points are during the articulation of the vowel [a]

Fig. 8 Location of region of interest (ROI) under the nose and the control area (C) where no turbulence was observed
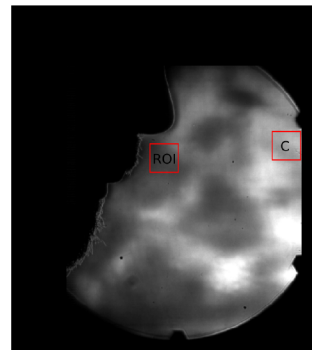
we expect anticipatory and overlay airflow in the [a] vowels, as has been demonstrated before for vowels preceding and following nasals (Beddor, 2015).

For the analysis, recordings were downsampled to 1 kHz. The warm air flowing out of the nose changes the density in the air that is mirrored by darker shades in the Schlieren photography (Figs. 3, 4, 5, 6 and 7). Thus, in order to analyze to what degree turbulences in the air stream were created by speech production, each video frame was transformed into a csv file that represents pixel intensity by means of a number ranging between 0 (= black) and 255 (= white). From these csv files, the region of interest (ROI) under the nose and a control area without any turbulence were extracted (see Fig. 8). All MATLAB, R and Python code used to process the
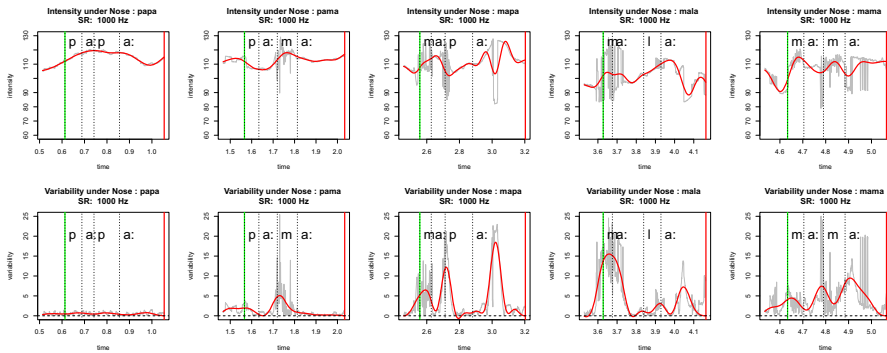
Fig. 9 Top row: Y-axis illustrates the time course (x-axis) of intensity (darkness of pixels) in different words in the region of interest under the nose in Figure 8. Solid vertical lines represent onset/offset of word; dashed vertical lines represent segment boundaries. Grey 'noisy' lines represent raw data, red line represents the smooth. Bottom row: Y-axis illustrates the time course (x-axis) of variability in intensity

data, in addition to the csv file, are available as Supplementary Materials (https://osf.io/z69ud/).

To simplify the analysis, we calculated the average shade value in each region for each frame. To normalize for background heat, average shade values in the control region were subtracted from the average shade values in the region of interest under the nose. For the analysis, normalized values were inverted such that high values represent darker shades and low values represent lighter shades (from now on called *intensity*). We analyzed the time course of two measures depending on nasal and non-nasal phones: intensity and the amount of variability in intensity under the nose. Stronger airflow out of the nose should be reflected by higher intensity and higher variability (due to fast changes in intensity in the region of interest).

Figure 9, top row, shows the time course of intensity (y-axis) in each of the five words. Green vertical lines indicate onset, red vertical lines indicate the offset of the word. The dotted vertical lines represent segment boundaries. The gray curve represents raw values, the red curve represents smoothed values. We will first focus on the smoothed curve.

As can be seen, there is little variation in intensity in 'papa'. In 'pama', intensity is lower in [p] and [a] than in [m]. In 'mapa' and 'mala', [m] and the following [a] show high intensity. In the same vein, the vowels in 'mama' show high intensity. However, the intensity measure is not conclusive, as it is also high across the entire 'papa' word. A more informative measure about airflow under the nose is the degree of *variability* in the region of interest as illustrated in the bottom row of Fig. 9. Variability (gray lines) represents the absolute difference between the raw intensity and the smoothed intensity in the top row in Fig. 9. The red line represents the smoothed variability.

In 'papa', only little variability can be observed, thus no airflow out of the nose. In 'pama', there is high variability at the offset of the pre-nasal [a] and during [m]. In 'mapa', variability starts before the onset of [m] and lasts until the onset of [a]. This indicates a delayed closing gesture of the velum which nasalizes the post-nasal

vowel. Interestingly, there is turbulence variability under the nose before [p]. This phenomenon has been described before as *nasal leakage*, which typically is related to vocal fold vibration before stops (Solé, 2011, 2018).

We also observe high variability in the center of the second [a] in 'mapa'. Given the raw intensity values, this could be an artifact. Turning our attention to 'mala', we see strong turbulence variability under the nose before [m] that lasts until the center of the first [a], showing an open velum during the vowel due to strong nasalization of the post-nasal vowel. Interestingly, whereas the first [a:] in 'mama' shows little turbulence variability, the second is strongly nasalized, as indicated by high variability.

We presented a qualitative analysis of one speaker uttering five words. The method can easily used to create data of multiple speakers uttering multiple words which can be subjected to linear and non-linear regression analyses such as mixed-effects regression (Bates et al., 2014) or generalized additive models (Wood, 2006). Regarding the frame rate, the presented analysis is based on data with a frame rate with 1000 Hz. We found that we obtain similar results with a frame rate of 500 Hz. Down-sampling the data further to 250 Hz and 125 Hz produced uninformative results (the plots for the different frame rate can be found in the supplementary materials). Thus, it seems that for the phenomenon at hand, a camera with at least 500 Hz frame rate would have been sufficient. In the next section we provide some guidelines for a Schlieren set up with a frame rate of 500 Hz.

3.3 Setting up a Schlieren system

Above we found that a frame rate of 10 kHz might not be necessary to detect nasalization. In fact a much more simplistic approach may be possible: it was shown e.g. by Strauss et al. (2019) that for Schlieren and Background Oriented Schlieren (BOS) setups with much more affordable equipment is sufficient even for measurements of supersonic flow. The recommended setup for lab experiments consists of a machine-vision industrial camera using e.g. the Sony® CMOS Pregius Sensor class (IMX287LLR) with a b/w frame rate up to 590 fps (see TheImagingSource, 2021).

Industrial cameras often feature C-mount lens mounts with a compatible standard camera lenses e.g. Zeiss® Planar or Nikon® lenses. The lens does not have to be a very special one, the only requirements are a sufficient focus length for the application (mostly zoom lenses) and a reasonable field of view to completely cover the CMOS sensor of the camera. The knife is a commonly available razor spare blade. The mentioned CREE® LED with a power of 3W and a colour temperature of 8000K is a simple and affordable point light source which does not need any type of aperture to turn it into a point light source (see also Settles, 2001, for details on how to set up a simple Schlieren setup).

The most expensive part of a simplified setup would be, besides the camera, a parabolic mirror of reasonable good optical quality. Flat mirrors do not work in a Schlieren setup. A possible choice would be the parabolic mirror range of Edmund Optics® (see EdmundOptics, 2021). Depending on the focus length such a mirror costs between 500 and 3000 Euro. If the alignment needs be facilitated even more,

optical holders, an optical bench or aluminium profile is recommended as well as a laser pointer or alignment laser in order to align the components on a repeatable basis. For the simplified setup only one mirror is needed (see Settles, 2001).

We do not recommend to use Newton reflector telescopes as a replacement. They often feature an alignment mark etched on their surface in the center that helps to align the main mirror with the secondary mirror. This mark will disrupt the image in the considered application.

All in all the costs for such a lab setup potentially range between 1500 and 6000 Euro, depending on the specific goals and on the needed field of view. Note that this setup is based on phenomenon at hand. We therefore recommend that readers who are interested in purchasing a Schlieren setup use the material provided in this corpus to test what frame rate and thus what camera is necessary to investigate the phenomenon they are interested in.

4 Conclusion

In the present paper, we have presented a corpus of Schlierenphotography, which allows to record air-dynamics in a non-intrusive manner. We have demonstrated potential ways to analyze the data and how to set up a system in the own lab.

References

Basset, P., Amelot, A., Vaissiére, J., & Roubeau, B. (2001). Nasal airflow in french spontaneous speech. *Journal of the International Phonetic Association, 31*(1), 87–99.
Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed- effects models using eigen and s4 [Computer software manual]. Retrieved Jan 2020 from http://CRAN.R-project.org/package=lme4
Beddor, P. (2015). The relation between language users' perception and production repertoires. In *Proceedings of the 18th international congress of phonetic sciences* (icphs iixx, glasgow).
Davidson, L. (2006). Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *Journal of the Acoustical Association of America, 120*(1), 407–415.

Davies, T. (1979). Schlieren photography: A tool for speech research. *Acoustics Letters, 3,* 73–75.

Davies, T. (1981). Schlieren photography short bibliography and re-view. *Optics & Laser Technology, 13*(1), 37–42. https://doi.org/10.1016/0030-3992(81)90089-X.

EdmundOptics, (2021). *Precision parabolic mirrors.* Retrieved Jan 2020 from https://www.edmundoptics.com/f/precision-parabolic-mirrors/11895

Hertegård, S., & Gauffin, J. (1992). Acoustic properties of the rothenberg mask. *STL- QPSR, 33*(2–3), 9–18.

Hoole, P., Mooshammer, C., & Tillmann, H. (1994). Kinematic analysis of vowel production in german. In *Proceedings of icslp 94, yokohama* (pp. 53-56).

Krane, M., & Gary, S. (2004). Aeroacoustics production of fricative speech sounds. *The Journal of the Acoustical Society of America, 115*(5), 2633–2633.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages* (Vol. 1012). Blackwell Oxford.

Mathiak, K., Klose, U., Ackermann, H., Hertrich, I., Kincses, W.-E., & Grodd, W. (2000). Stroboscopic articulography using fast magnetic resonance imaging. *International Journal of Language & Communication Disorders, 35*(3), 419–425. https://doi.org/10.1080/136828200410663.

Mooshammer, C., & Fuchs, S. (2002). Stress distinction in german: Simulating kinematic parameters of tongue-tip gestures. *Journal of Phonetics, 30*(3), 337–355.

Petrone, C., Fuchs, S., & Koenig, L. L. (2017). Relations among subglottal pressure, breathing, and acoustic parameters of sentence-level prominence in german. *The Journal of the Acoustical Society of America, 141*(3), 1715–1725.

Quigley, J., Shiere, F., Webster, R. C., & Cobb, C. M. (1964). Measuring palatopharyngeal comptence with the nasal anemometer. *The Cleft Palate Journal, 1*(3), 304–313.

Rowell, J., Noguchi, M., Bernhardt, B. M., Herdman, A., Gick, B., & Schellenberg, M. (2016). Schlieren study of external airflow during the production of nasal and oral vowels in french. *Canadian Acoustics, 44*(3).

Settles, G. (2001). *Schlieren and shadowgraph techniques* (1st ed.). Springer.

Solé, M.-J. (2011). Voice-initiating Gestures in Spanish: Prenasalization. UC Berkeley PhonLab Annual Report, *7*(7). Retrieved December 10, 2020, from https://escholarship.org/uc/item/0p25r1pv

Solé, M.-J. (2018). Articulatory adjustments in initial voiced stops in Span-ish, French and English. *Journal of Phonetics, 66,* 217–241. https://doi.org/10.1016/j.wocn.2017.10.002.

Strauss, F. T., General, S., Manfletti, C., & Schlechtriem, S. (2019). Flow visualizations with back-ground oriented Schlieren in a transpiration-cooled model scramjet combustor. *International Journal of Energetic Materials and Chemical Propulsion, 18*(2), 133–155. https://doi.org/10.1615/IntJEnergeticMaterialsChemProp.2019028003.

Strauss, F. T., Man etti, C., Freudenmann, D., Witte, J., & Schlechtriem, S. (2016). Preliminary experiments on transpiration cooling in ramjets and scramjets. In *52nd aiaa/sae/asee joint propulsion conference* (p. AIAA2016-4968).

Strauss, F. T., Witte, J., Weisswange, M., Manfletti, C., & Schlechtriem, S. (2017). Experiments on shock-boundary layer interaction and cooling efficiency in a transpiration cooled model scramjet. In *53rd aiaa/sae/asee joint propulsion conference* (p. AIAA2017-4833). 15

TheImagingSource. (2021). Industrial cameras. Retrieved February 27, 2021, from https://www.theimagingsource.com/products/industrial-cameras/usb-3.0-monochrome/dmk33ux287/

Tiede, M. K., Perkell, J., Zandipour, M., & Matthies, M. (2001). Gestural timing effects in the "perfect memory" sequence observed under three rates by electromagnetometry. *The Journal of the Acoustical Society of America, 110*(5), 2657–2657. https://doi.org/10.1121/1.4777046.

Tomaschek, F., Tucker, B. V., Baayen, R. H., & Fasiolo, M. (2018). Practice makes perfect: The consequences of lexical proficiency for articulation. *Linguistic Vanguard, 4*(s2), 1–13.

Uecker, M., Zhang, S., Voit, D., Karaus, A., Merboldt, K.-D., & Frahm, J. (2010). Real-time mri at a resolution of 20 ms. *NMR in Biomedicine, 23*(8), 986–994.

Van Dyke, M. (1997). *An album of uid motion (5th* (edition). The Parabolic Press.

Warren, D. W. (1967). Nasal emission of air and velopharyngeal function. *The Cleft Palate Journal, 4*(2), 148–156.

Wood, S. N. (2006). *Generalized additive models: An introduction with r.* Chapman and Hall/CRC.

Wrench, A. A., & Scobbie, J. M. (2011). Very high frame rate ultrasound tongue imaging. In *Proceedings of the 9th international seminar on speech production (issp)* (pp. 155–162).

Zharkova, N., Hewlett, N., & Hardcastle, W. J. (2012). An ultrasound study of lingual coarticulation in /
      sv/ syllables produced by adults and typically developing children. *Journal of the International
      Phonetic Association, 42*(2), 193–208.

Publisher's Note   Springer Nature remains neutral with regard to jurisdictional claims in published
maps and institutional affiliations.